ABHILASH SRIVASTAVA

LinkedIn — GitHub

Email: abhilash.shakti@gmail.com
Mobile: 858-349-7959

EXPERIENCE

Apollo.io — Remote
Senior ML Engineer — *Aug 2023 - Present*

- Led the ML development and launch of "Ask Apollo" and "Pre-meeting Insights" in Q3 '24, utilizing the latest GenAI techniques (Gemini and OpenAI GPT models) to build a first-in-class Sales Copilot.
- Drove the successful launch of three advanced AutoScoring features (Keywords, Technologies, Job Titles) in Q1 '24, fueling a 128% surge in Weekly Active Users and a 33% uplift in search engagement across 12,000+ teams.
- Enhanced Call Summary insights through targeted model improvements, achieving a 22% increase in precision and 14% in recall, and spearheaded CI innovations like follow-up emails and "Ask Anything V1", showcasing technical depth and customer impact.
- Led a complex online metrics experiment to prioritize AutoScoring for GA, collaborating with Product Analytics and Prospecting teams to overcome metrics throttling, streamline iterations, and drive data-informed decisions.

Meta — Bellevue, WA
Senior ML Engineer — *Apr 2022 - May 2023*

- Developed and deployed advanced ML/DL models with continuous learning to detect fraud, cloaking, and malicious ads, blocking 50k+ malicious accounts daily and saving Facebook $70M monthly.
- Enhanced ML model transparency and governance using multi-class/multi-label classifiers for 100+ policy violations, reducing false-discovery rates via large-scale user-engagement analysis.
- Built and optimized GB-scale Hive data pipelines for training data and human labeling, collaborating with infrastructure teams to accelerate model training and boost developer efficiency.
- Mentored new hires, led candidate interviews, supported user privacy, participated in cross-functional initiatives, on-call rotations, and roadmap planning.

Microsoft — Redmond, WA
Software Engineer — *May 2019 - Mar 2022*

- Enhanced Enterprise Search with query understanding and ranker models, multilingual ML pipeline (data, modeling, deployment, feedback) serving 13k+ organizations across 20M+ monthly queries in 15+ languages.
- Developed multilingual and query language identification models (BERT/XLM-R, student-teacher distillation), differentiating over 100 languages and powering Bing and Office products internationally.
- Scaled a unified language model for global domain classification, intent detection, and slot tagging, expanding support across multiple GEOs through Universal Encoder/Unicode fine-tuning.
- Mentored 2 Ph.D. interns (Information Retrieval/NLP) and contributed as a reviewer for MLADS and as a Microsoft AI for Accessibility (AI4A) grants committee member.

TECH SKILLS

- **Languages:** Python, JAVA, JavaScript, C#, SQL
- **ML/DL Tech:** LLMs, RAG, Pytorch, Keras, TensorFlow, PySpark
- **Data Engineering Tech:** GCP, VertexAI, Azure, AzureML, Databricks, Snowflake, Hadoop, Airflow, Amplitude
- **Other:** Analytics, Humanloop, Dashboards, Metrics, Legal/Regulatory Guidelines

EDUCATION

University of California, San Diego — San Diego, CA
M.S. in Computer Science (Artificial Intelligence) — *Sep 2017 – Mar 2019*

Birla Institute of Technology and Science, Pilani — Pilani, India
B.E. in Computer Science — *Aug 2010 – Jun 2014*

PATENTS

- **Method and system of sharing vehicle health information among vehicles (US20180218609A1):** Identifying the suitability of driving a vehicle in a specific region using and sharing vehicular data in vicinity.
- **Methods and systems for rendering user intent across electronic devices (US20180219988A1):** Intent prediction based on user action history, current application context and trigger from smart input device.